Research Article   **Open Access**

Michele Benzi* and Ru Huang

Some matrix properties preserved by generalized matrix functions

Abstract: Generalized matrix functions were first introduced in [J. B. Hawkins and A. Ben-Israel, *Linear and Multilinear Algebra*, 1(2), 1973, pp. 163–171]. Recently, it has been recognized that these matrix functions arise in a number of applications, and various numerical methods have been proposed for their computation. The exploitation of structural properties, when present, can lead to more efficient and accurate algorithms. The main goal of this paper is to identify structural properties of matrices which are preserved by generalized matrix functions. In cases where a given property is not preserved in general, we provide conditions on the underlying scalar function under which the property of interest will be preserved by the corresponding generalized matrix function.

1 Introduction

A generalized matrix function (GMF) is a type of matrix function that is defined in terms of the singular value decomposition (SVD), and that can be applied to rectangular matrices. For many years after their introduction in [9], there was very little interest in GMFs. Recently, however, it has been recognized that GMFs provide a framework for formulating and solving various problems arising in data science, ranging from rank-constrained matrix optimization problems to the analysis of directed networks; see, e.g., [1, 2] and the references therein. GMFs also arise naturally in the context of Hamiltonian dynamical systems [7, 8], such as the wave equation on graphs [5]. This realization has prompted several authors to take a new look at GMFs, leading to several papers, both theoretical and computational in nature [1, 3, 5, 15].

This paper is aimed at deepening our theoretical understanding of generalized matrix functions, in particular we are interested in studying those matrix properties that are invariant under generalized matrix functions. In the case of standard matrix functions, advance knowledge of the structural properties of $f(A)$ can lead to more accurate and efficient algorithms; for example, when A is a triangular Toeplitz matrix and f is a function such that $f(A)$ is defined, the fact that $f(A)$ is also triangular and Toeplitz can lead to significant savings when computing it. Similar savings may be expected in the case of generalized matrix functions when the matrix A has certain structural properties. Some of the properties preserved by standard matrix functions are not preserved by generalized matrix functions; conversely, there exist structural properties preserved by generalized matrix functions which are not preserved by standard matrix functions. It is therefore necessary to systematically investigate the structural properties that are preserved by GMFs and, more generally, the interplay between matrix structures and GMFs.

*Corresponding Author: Michele Benzi:** Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy,
E-mail: michele.benzi@sns.it
Ru Huang: Department of Mathematics, Emory University, Atlanta, Georgia 30322, USA, E-mail: ru.huang@emory.edu.

The remainder of the paper is organized as follows. In section 2 we give some basic definitions and review useful properties of GMFs. In section 3 we prove the invariance under GMFs of a number of different matrix properties, and we derive a few invariance results under some conditions on the scalar function f for other types of matrix structures. In section 4 we show that GMFs are well behaved with respect to the isomorphism between complex $n \times n$ matrices and their canonical real representation as $2n \times 2n$ real matrices. A few additional results are given in section 5, where we consider the invariance of certain matrix cones under GMFs. Section 6 contains a few concluding remarks and points to possible topics for further research.

2 Preliminaries

Let $A \in \mathbb{C}^{m \times n}$. Consider the singular value decomposition

$$A = U\Sigma V^H,$$

with U and V unitary and Σ being a diagonal matrix with the singular values of A on the main diagonal in non-increasing order [13]. Let r be the rank of A. Consider the matrices U_r and V_r formed with the first r columns of U and V, and let Σ_r be the leading $r \times r$ principal submatrix of Σ whose diagonal entries are $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. Then we have the compact SVD:

$$A = U_r \Sigma_r V_r^H.$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function such that $f(\sigma_i)$ is defined for all $i = 1, 2, \ldots, r$. Define the generalized matrix function induced by f as

$$f^\diamond(A) = U_r f(\Sigma_r) V_r^H, \tag{1}$$

where

$$f(\Sigma_r) = \begin{bmatrix} f(\sigma_1) & & & \\ & f(\sigma_2) & & \\ & & \ddots & \\ & & & f(\sigma_r) \end{bmatrix}.$$

We observe that $f^\diamond(A)$ reduces to the standard matrix function $f(A)$ whenever A is Hermitian positive definite, or when A is Hermitian positive semidefinite and f satisfies $f(0) = 0$. We also note that the rank of $f^\diamond(A)$ can never exceed r, the rank of A, a property which does not hold, in general, for standard matrix functions.

Another important observation is that we can assume f to be an odd function, since the function f need only be defined for positive real values, and any such function can be completed to an odd function by setting $f(0) = 0$ and $f(-x) = -f(x)$ for all $x < 0$ (note that the values of f outside the interval $[-\sigma_1, \sigma_1]$ are irrelevant). Setting

$$\mathscr{A} = \begin{bmatrix} 0 & A \\ A^H & 0 \end{bmatrix}, \tag{2}$$

we have that any real-valued odd function defined on $[-\sigma_1, \sigma_1]$ satisfies

$$f(\mathscr{A}) = \begin{bmatrix} 0 & f^\diamond(A) \\ f^\diamond(A)^H & 0 \end{bmatrix}. \tag{3}$$

Equation (3) is one of several important formulas relating generalized matrix functions with standard ones. Additional basic properties of generalized matrix functions are given in the following propositions.

Proposition 1. *([3]) Let $A \in \mathbb{C}^{m \times n}$ be a matrix of rank r. Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function, and let $f^\diamond : \mathbb{C}^{m \times n} \to \mathbb{C}^{m \times n}$ be the induced generalized matrix function. Then*

(i) $[f^\diamond(A)]^H = f^\diamond(A^H)$.

(ii) Let $X \in \mathbb{C}^{m \times m}$ and $Y \in \mathbb{C}^{n \times n}$ be two unitary matrices, then $f^\diamond(XAY) = X[f^\diamond(A)]Y$.

(iii) If $A = A_1 \oplus A_2 \oplus \cdots \oplus A_k$, then $f^\diamond(A) = f^\diamond(A_1) \oplus f^\diamond(A_2) \oplus \cdots \oplus f^\diamond(A_k)$.

(iv) $f^\diamond(A) = f(\sqrt{AA^H})(\sqrt{AA^H})^\dagger A = A(\sqrt{A^H A})^\dagger f(\sqrt{A^H A})$, where $M^\dagger$ is the Moore–Penrose pseudoinverse of M.

We will also make use of the following simple result.

Proposition 2. *Let $A \in \mathbb{C}^{m \times n}$ be a matrix of rank r. Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function, and let $f^\diamond : \mathbb{C}^{m \times n} \to \mathbb{C}^{m \times n}$ be the induced generalized matrix function. Then*

(i) $[f^\diamond(A)]^T = f^\diamond(A^T)$.

(ii) $\overline{f^\diamond(A)} = f^\diamond(\bar{A})$.

Proof. (i) Assume A has the SVD: $A = U\Sigma V^H$. Observing that U^T and $\bar{V}$ are also unitary, A^T has the compact SVD: $A^T = \bar{V}_r \Sigma_r U_r^T$, hence $f^\diamond(A^T) = \bar{V}_r f(\Sigma_r) U_r^T = f^\diamond(A)^T$.

(ii) The proof is similar to (i) . □

3 Matrix structures preserved by GMFs

In this section we identify several classes of matrices whose properties are preserved under (certain) generalized matrix functions. We deal primarily (but not exclusively) with square matrices, since there are relatively few types of structured rectangular (non-square) matrices. We emphasize that since $f^\diamond(A)$ cannot (in general) be represented as a polynomial in A, it is not typically true that associative matrix algebras are closed under GMFs. This is a major difference with respect to standard matrix functions. Nevertheless, we will see that a number of important matrix classes are indeed closed under GMFs.

Similar to [10, page 314], we present several important matrix structures in Table 1. These matrices correspond to the classical Lie and Jordan matrix algebras over the real and the complex numbers, and can be described in terms of symmetry, or anti-symmetry, with respect to a prescribed bilinear or sesquilinear form. In the Table, the following notation is used:

$$R_n = \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}, \quad J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \quad \Sigma_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \quad \text{with} \quad p + q = n,$$

and

$$A^\star := \begin{cases} M^{-1}A^T M & \text{for bilinear forms,} \\ M^{-1}A^H M & \text{for sesquilinear forms,} \end{cases}$$

where M is one of the matrices defining the above bilinear or sesquilinear forms. We remark that we adhere here to the terminology used in [10], and that different terms are often found in the literature for some of the matrix structures discussed below. Our first result is the following.

Theorem 3. *Let $\mathcal{M}$ be one of the classes in Table 1. If $A \in \mathcal{M}$ and $f^\diamond(A)$ is well defined, then $f^\diamond(A) \in \mathcal{M}$.*

Proof. It is obvious that R_n, J and $\Sigma_{p,q}$ are unitary, so in each case the matrix M corresponding to the algebra to which A belongs is unitary. Thus by Propositions 1(ii) and 2 we have

$$f^\diamond(A^\star) = \begin{cases} M^{-1}f^\diamond(A)^T M = f^\diamond(A)^\star & \text{for bilinear forms,} \\ M^{-1}f^\diamond(A)^H M = f^\diamond(A)^\star & \text{for sesquilinear forms.} \end{cases}$$

Hence, for the Jordan algebra case, we have $f^\diamond(A)^\star = f^\diamond(A^\star) = f^\diamond(A)$, and similarly for the Lie algebra case we have $f^\diamond(A)^\star = f^\diamond(A^\star) = f^\diamond(-A) = -f^\diamond(A)$. □

Table 1: Structured matrices associated with certain bilinear and sesquilinear forms

Space	M	Jordan Algebra $\mathscr{J} = \{A : A^* = A\}$	Lie Algebra $\mathscr{L} = \{A : A^* = -A\}$
		Bilinear forms	
$\mathbb{R}^n$	I	Symmetrics	Skew-symmetrics
$\mathbb{C}^n$	I	Complex symmetrics	Complex skew-symmetrics
$\mathbb{R}^n$	$\Sigma_{p,q}$	Pseudosymmetrics	Pseudoskew-symmetrics
$\mathbb{C}^n$	$\Sigma_{p,q}$	Complex pseudo-symmetrics	Complex pseudo-skew-symmetrics
$\mathbb{R}^n$	R_n	Persymmetrics	Perskew-symmetrics
$\mathbb{R}^{2n}$	J	Skew-Hamiltonians	Hamiltonians
$\mathbb{C}^{2n}$	J	Complex J-skew-symmetrics	Complex J-symmetrics
		Sesquilinear forms	
$\mathbb{C}^n$	I	Hermitians	Skew-Hermitians
$\mathbb{C}^n$	$\Sigma_{p,q}$	Pseudo-Hermitians	Pseudoskew-Hermitians
$\mathbb{C}^n$	R_n	Perhermitians	Skew-perhermitians
$\mathbb{C}^{2n}$	J	J-skew-Hermitians	J-Hermitians

A simple but worth mentioning consequence of this result is the observation that for any matrix A of rank r, the structure of the left and right singular vectors u_i, v_i ($1 \le i \le r$) alone determines whether or not A belongs to one of the matrix classes in Table 1. The singular values $\sigma_1, \ldots, \sigma_r$ play no role, since replacing them with arbitrary real numbers produces a matrix in the same class.

There are several other matrix classes that are preserved by arbitrary generalized matrix functions. We recall the definition of a centrohermitian (skew-centrohermitian) matrix, not necessarily square [14].

Definition 4. $A \in \mathbb{C}^{m \times n}$ is centrohermitian (skew-centrohermitian) if $R_m A R_n = \bar{A}$ (respectively, $R_m A R_n = -\bar{A}$).

The next result provides another list of structured matrix classes that are preserved by any well-defined generalized matrix function.

Theorem 5. *Let $A \in \mathbb{C}^{m \times n}$ be a matrix of rank r. Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function, and let $f^\diamond : \mathbb{C}^{m \times n} \to \mathbb{C}^{m \times n}$ be the induced generalized matrix function, which is assumed to be well defined at A.*

(i) If A is centrohermitian (skew-centrohermitian), then $f^\diamond(A)$ is also centrohermitian (skew-centrohermitian).

(ii) If $m = n$ and A is normal, then $f^\diamond(A)$ is also normal.

(iii) If $m = n$ and A is circulant, then $f^\diamond(A)$ is also circulant.

(iv) If A is a block-circulant matrix with circulant blocks, then $f^\diamond(A)$ is also a block-circulant matrix with circulant blocks. This generalizes (iii).

Proof. (i) First, if A is centrohermitian, then $R_m A R_n = \bar{A}$. Since R_m and R_n are unitary, by Propositions 1 (ii) and 2 (ii) we have $R_m f^\diamond(A) R_n = f^\diamond(R_m A R_n) = f^\diamond(\bar{A}) = \overline{f^\diamond(A)}$.

Second, if A is skew-centrohermitian, then $R_m A R_n = -\bar{A}$. Hence, $R_m f^\diamond(A) R_n = f^\diamond(R_m A R_n) = f^\diamond(-\bar{A}) = -f^\diamond(\bar{A}) = -\overline{f^\diamond(A)}$.

(ii) Since A is normal, there exists $V \in \mathbb{C}^{n \times n}$ unitary such that $A = V \Lambda V^H$ where Λ is diagonal with the eigenvalues λ_i of A on the main diagonal. Assuming that A has rank r and that the eigenvalues of A are numbered so that $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_r| > |\lambda_{r+1}| = \cdots = |\lambda_n| = 0$, we can write $A = V Q \Sigma V^H$ where

$$Q = \mathrm{diag}\left(\frac{\lambda_1}{|\lambda_1|}, \ldots, \frac{\lambda_r}{|\lambda_r|}, 1, \ldots, 1\right) \quad \text{and} \quad \Sigma = \mathrm{diag}(|\lambda_1|, \ldots, |\lambda_r|, 0, \ldots, 0).$$

Noting that Q is unitary, A has the singular value decomposition $A = U\Sigma V^H$ with $U = VQ$. Let $A = U_r \Sigma_r V_r^H$ be the corresponding compact SVD, then

$$f^\diamond(A) = U_r f(\Sigma_r) V^H = VDV^H$$

with

$$D = \text{diag}\left(\frac{\lambda_1}{|\lambda_1|} f(|\lambda_1|), \ldots, \frac{\lambda_r}{|\lambda_r|} f(|\lambda_r|), 0, \ldots, 0 \right). \tag{4}$$

Therefore, $f^\diamond(A)$ is normal.

(iii) Recall that circulant matrices are precisely those that are diagonalized by the Discrete Fourier Transform (DFT) matrix

$$U_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \ldots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2n-2} & \ldots & \omega^{(n-1)^2} \end{bmatrix},$$

where $\omega = e^{-\frac{2\pi i}{n}}$; see [6, Theorems 3.3.2–3.3.3]. Thus, we have

$$A = U\Lambda U^H,$$

where Λ is diagonal and contains the eigenvalues of A. Arguing as in the proof of (ii), we conclude that $f^\diamond(A) = UDU^H$ where D is given by (4), hence $f^\diamond(A)$ is diagonalized by the DFT matrix and is therefore circulant.

(iv) Let $U = U_m \otimes U_n$, with U_m and U_n being the DFT matrices of dimension m and n, respectively. Note that U is unitary. It is known that $A \in \mathbb{C}^{mn \times mn}$ is a block-circulant matrix with $m \times m$ blocks, where each block is a $n \times n$ circulant matrix, if and only if A has the decomposition $A = U\Lambda U^*$ where Λ is a diagonal matrix [6, Theorem 5.8.1]. Then using a similar argument as in (iii) we obtain the result. $\qquad \square$

Remark 6. *We should mention that in [9], Hawkins and Ben-Israel adopt a somewhat different definition of generalized matrix function which allows for "complex singular values"; i.e., the (real) singular values σ_j of A can be multiplied by (complex) phase factors $e^{i\theta_j}$, where $\theta_j \in [0, 2\pi)$. For $f^\diamond(A)$ to be defined, the scalar function f needs to be defined at all these possibly complex values. It is clear that with this definition of $f^\diamond$, the results in of items (ii)-(iv) would be immediate. When using our definition of $f^\diamond$, which only allows for real singular values and functions f of a real variable, the results are not immediately obvious and need proof.*

While normality, as we saw, is always preserved by GMFs, the weaker property of being diagonalizable is not, generally speaking. As an example, let $\sigma_1 > \sigma_2 > 0$ and consider the 2×2 matrix

$$A = \begin{bmatrix} 0 & \sigma_1 \\ \sigma_2 & 0 \end{bmatrix} = \sigma_1 e_1 e_2^T + \sigma_2 e_2 e_1^T.$$

Note that A is diagonalizable (having two distinct eigenvalues). If f is any function defined for $x > 0$ such that $f(\sigma_1) \neq 0$ and $f(\sigma_2) = 0$ (or $f(\sigma_1) = 0$ and $f(\sigma_2) \neq 0$), the matrix $f^\diamond(A)$ is not diagonalizable. Note that this cannot happen with standard matrix functions.

The next result concerns the preservation of zero entries in certain positions by generalized matrix functions.

Theorem 7. *Let $A \in \mathbb{C}^{m \times n}$ and let $f^\diamond(A)$ be defined.*

(i) If the ith column (row) of A consists of all zeros, then the ith column (row) of $f^\diamond(A)$ also consists of all zeros.

(ii) If there exist permutation matrices $P \in \mathbb{R}^{n \times n}$ and $Q \in \mathbb{R}^{n \times n}$ such that PAQ is block diagonal, then the same property holds for $f^\diamond(A)$.

Proof. (i) Without loss of generality, we may assume the last column of A is 0, since $f^\diamond(AP) = f^\diamond(A)P$ for any permutation matrix P. Write $A = \begin{bmatrix} \hat{A} & 0 \end{bmatrix}$ and assume we have an SVD of $\hat{A}$,

$$\hat{A} = \hat{U}\hat{\Sigma}\hat{V}^H.$$

This induces an SVD of A:

$$A = \begin{bmatrix} \hat{A} & 0 \end{bmatrix} = \hat{U} \begin{bmatrix} \hat{\Sigma} & 0 \end{bmatrix} \begin{bmatrix} \hat{V}^H & 0 \\ 0 & 1 \end{bmatrix} = U\Sigma V^H.$$

Assuming A has rank r, we have that

$$f^\diamond(A) = U_r f(\Sigma_r) V_r^H$$

with U_r, Σ_r and V_r defined as usual. Observing that the last row of V_r consists of all zeros, we conclude that the last column of $f^\diamond(A)$ is also 0.

Similarly, for the case that A has a row consisting of all zeros, we have that A^H has a column of all zeros. Using the fact that $f^\diamond(A)^H = f^\diamond(A^H)$ we easily obtain the result.

(ii) This is a straightforward consequence of Proposition 1 (ii)-(iii). We emphasize that A need not be square, and neither do the individual blocks. □

For some structured matrices, although their structural properties may not be preserved by every generalized matrix function, we can place some restrictions on the function f so that $f^\diamond(A)$ has the same structure as A. We first illustrate this on matrix groups corresponding to the Lie and Jordan algebras in Table 1. Our classification closely follows [10, page 314]. Here the matrices R, J, $\Sigma_{p,q}$ and $A^\star$ are defined as before. We have the following simple result.

Theorem 8. *Let $\mathscr{G}$ be one of the matrix groups in Table 2. If $A \in \mathscr{G}$, f is defined for $x > 0$ and satisfies $f(x)f(\frac{1}{x}) = 1$ for $x > 0$, then $f^\diamond(A) \in \mathscr{G}$.*

Proof. Since R_n, J and $\Sigma_{p,q}$ are unitary, so is the matrix M in each case corresponding to the second column of Table 2. Thus by Proposition 1 and 2 we have

$$f^\diamond(A^\star) = \begin{cases} M^{-1}f^\diamond(A)^T M = f^\diamond(A)^\star & \text{for bilinear forms,} \\ M^{-1}f^\diamond(A)^H M = f^\diamond(A)^\star & \text{for sesquilinear forms.} \end{cases}$$

Hence, for A in each of the above matrix groups, we have $f^\diamond(A)^\star = f^\diamond(A^\star) = f^\diamond(A^{-1}) = V_r f(\Sigma_r^{-1})U_r^H = V_r f(\Sigma_r)^{-1} U_r^H = f^\diamond(A)^{-1}$. □

We remark that the condition $f(x)f(x^{-1}) = 1$ (for all $x > 0$) is satisfied in particular by the functions $f(x) = x^\alpha$ and $f(x) = -x^\alpha$ with $\alpha \in \mathbb{R}$ arbitrary. A complete characterization of all (meromorphic) functions satisfying such condition can be found in [11]. We also note the recent work [4], where the general problem of determining conditions on f that guarantee the preservation of a given matrix structure by the corresponding $f^\diamond$ is briefly mentioned.

We note that in the case of a unitary matrix A we have $f^\diamond(A) = f(1)A$, therefore $f^\diamond(A) = A$ for any scalar function f satisfying $f(1) = 1$. Hence, generalized matrix functions behave trivially on unitary matrices. This holds more generally for any matrix of the form $A = U_r V_r^H$ where U_r, $V_r \in \mathbb{C}^{n \times r}$ have orthonormal columns.

The next result concerns preservation of nonnegativity of the entries. Recall that if f is an analytic function expressed by a Maclaurin series expansion with nonnegative coefficients and A is a square nonnegative matrix, then the standard matrix function $f(A)$, if defined, is also nonnegative. The corresponding result for generalized matrix functions of possibly rectangular nonnegative matrices is the first part of the following theorem.

Theorem 9. *(i) Let $A \in \mathbb{R}^{m \times n}$ be nonnegative and let f be the odd part of an analytic function which has the Maclaurin expansion of the form $f(z) = \sum_{k=0}^{\infty} c_k z^k$ with $c_{2k+1} \geq 0$, assumed to be convergent for $|z| < R$ with $R > \|A\|_2$. Then $f^\diamond(A)$ is well-defined, and $f^\diamond(A)$ is also nonnegative.*

Table 2: Matrix groups associated with certain bilinear and sesquilinear forms

Space	M	Automorphism Group $\mathscr{G} = \{A : A^{\star} = A^{-1}\}$
		Bilinear forms
$\mathbb{R}^n$	I	Real orthogonals
$\mathbb{C}^n$	I	Complex orthogonals
$\mathbb{R}^n$	$\Sigma_{p,q}$	Pseudo-orthogonals
$\mathbb{C}^n$	$\Sigma_{p,q}$	Complex pseudo-orthogonals
$\mathbb{R}^n$	R_n	Real perplectics
$\mathbb{R}^{2n}$	J	Real symplectics
$\mathbb{C}^{2n}$	J	Complex symplectics
		Sesquilinear forms
$\mathbb{C}^n$	I	Unitaries
$\mathbb{C}^n$	$\Sigma_{p,q}$	Pseudo-unitaries
$\mathbb{C}^n$	R_n	Complex perplectics
$\mathbb{C}^{2n}$	J	Conjugate symplectics

(ii) If $A \in \mathbb{R}^{n \times n}$ is doubly stochastic, f satisfies the same assumptions as in (i) and in addition $f(1) = 1$, then $f^{\diamond}(A)$ is also doubly stochastic.

Proof. (i) Without loss of generality, we can assume that f is odd and that $f(x) = \sum_{k=0}^{\infty} c_{2k+1} x^{2k+1}$, a series with nonnegative coefficients. The condition on the radius of convergence of the Maclaurin series guarantees that $f^{\diamond}(A)$ is well-defined. The result then follows immediately from (3), noting that the matrix $\mathscr{A}$ in (2) is nonnegative. Alternatively, assume we have the compact SVD of A: $A = U_r \Sigma_r V_r^T$, then $AA^T = U_r \Sigma_r^2 U_r^T \geq 0$. Therefore $(AA^T)^k A = U_r \Sigma_r^{2k+1} V_r^T \geq 0$. It follows that $f^{\diamond}(A) = U_r f(\Sigma_r) V_r^T = U_r (\sum_{k=0}^{\infty} c_{2k+1} \Sigma_r^{2k+1}) V_r^T = \sum_{k=0}^{\infty} c_{2k+1} U_r \Sigma_r^{2k+1} V_r^T \geq 0$.

(ii) First note that $f^{\diamond}(A)$ is nonnegative by (i). Next, let e be the vector of all 1. Then for a doubly stochastic matrix we have

$$Ae = A^T e = e.$$

Hence,

$$U_r \Sigma_r V_r^T e = e \qquad \text{and} \qquad V_r \Sigma_r U_r^T e = e.$$

Left-multiplying by U_r^T both sides of the first equation and by $\Sigma_r V_r^T$ both sides of the second equation, we get

$$\Sigma_r V_r^T e = U_r^T e \qquad \text{and} \qquad \Sigma_r^2 U_r^T e = \Sigma_r V_r^T e.$$

Combining the equations above we have

$$\Sigma_r^2 V_r^T e = V_r^T e \qquad \text{and} \qquad \Sigma_r^2 U_r^T e = U_r^T e.$$

Note that $A^T A$ is also doubly stochastic, hence the largest singular value of A is equal to 1. Let k be its multiplicity. Let $V_r^T e = \begin{bmatrix} b_1 & b_2 & \dots & b_r \end{bmatrix}^T$, then the first of the two equations above becomes

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \sigma_{k+1}^2 & & \\ & & & & \ddots & \\ & & & & & \sigma_r^2 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ b_{k+1} \\ \vdots \\ b_r \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ b_{k+1} \\ \vdots \\ b_r \end{bmatrix}.$$

Hence, $\sigma_i^2 b_i = b_i$, for $i = k + 1, \ldots, r$. Because $\sigma_i < 1$, we have $b_i = 0$, for $i = k + 1, \ldots, r$. Therefore,

$$f^\Diamond(A)e = U_r f(\Sigma_r) V_r^T e$$

$$= U_r \begin{bmatrix} f(1) & & & & & \\ & \ddots & & & & \\ & & f(1) & & & \\ & & & f(\sigma_{k+1}) & & \\ & & & & \ddots & \\ & & & & & f(\sigma_r) \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} = U_r \begin{bmatrix} f(1)b_1 \\ \vdots \\ f(1)b_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= f(1)U_r \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} = f(1)Ae = f(1)e = e.$$

Hence, $f^\Diamond(A)$ is row-stochastic. A similar argument can be used to show that $f^\Diamond(A)$ is also column-stochastic and thus doubly stochastic. □

4 GMFs and the complex-to-real isomorphism

In this section we show that GMFs are well-behaved with respect to the canonical isomorphism between the algebra of $n \times n$ complex matrices (over the reals) and the subalgebra of the algebra of real $2n \times 2n$ matrices consisting of all block matrices of the form $\begin{bmatrix} B & -C \\ C & B \end{bmatrix}$ where $B, C \in \mathbb{R}^{n \times n}$.

Theorem 10. *Let $A = B + iC \in \mathbb{C}^{n \times n}$ (with B and C real) have rank r. Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function; without loss of generality, we assume that $f(0) = 0$ if $r < n$. Let $f^\Diamond : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ be the induced generalized matrix function. Let $\Phi : \mathbb{C}^{n \times n} \to \mathbb{R}^{2n \times 2n}$ be the mapping $\Phi(A) = \begin{bmatrix} B & -C \\ C & B \end{bmatrix}$. We also denote by $f^\Diamond$ the generalized matrix function from $\mathbb{R}^{2n \times 2n}$ to $\mathbb{R}^{2n \times 2n}$ induced by f. Then $f^\Diamond(\Phi(A))$ is well defined and $f^\Diamond$ commutes with Φ:*

$$f^\Diamond(\Phi(A)) = \Phi(f^\Diamond(A)).$$

Proof. We have

$$\begin{aligned} A &= B + iC \\ &= U\Sigma V^H \\ &= (U_1 + iU_2)\Sigma(V_1 + iV_2)^H \\ &= (U_1 + iU_2)\Sigma(V_1^T - iV_2^T) \\ &= U_1\Sigma V_1^T + U_2\Sigma V_2^T + i(U_2\Sigma V_1^T - U_1\Sigma V_2^T), \end{aligned}$$

and $f^\Diamond(A) = U_1 f(\Sigma) V_1^T + U_2 f(\Sigma) V_2^T + i(U_2 f(\Sigma) V_1^T - U_1 f(\Sigma) V_2^T)$. Hence,

$$\Phi(f^\Diamond(A)) = \begin{bmatrix} U_1 f(\Sigma) V_1^T + U_2 f(\Sigma) V_2^T & -U_2 f(\Sigma) V_1^T + U_1 f(\Sigma) V_2^T \\ U_2 f(\Sigma) V_1^T - U_1 f(\Sigma) V_2^T & U_1 f(\Sigma) V_1^T + U_2 f(\Sigma) V_2^T \end{bmatrix}.$$

Next, observe that $\Phi(A) = \begin{bmatrix} B & -C \\ C & B \end{bmatrix}$ has the decomposition:

$$\begin{bmatrix} B & -C \\ C & B \end{bmatrix} = \begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} V_1 & -V_2 \\ V_2 & V_1 \end{bmatrix}^T$$

.

Because U is unitary, we have $(U_1 + iU_2)(U_1^T - iU_2^T) = I$ and $(U_1^T - iU_2^T)(U_1 + iU_2) = I$, therefore $U_1 U_1^T + U_2 U_2^T = I$ and $U_1 U_2^T = U_2 U_1^T$.

Thus $\begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix}$ is orthogonal. Similarly, $\begin{bmatrix} V_1 & -V_2 \\ V_2 & V_1 \end{bmatrix}$ is also orthogonal.

By Proposition 1, parts (ii) and (iii), we have

$$\begin{aligned} f^\diamond(\Phi(A)) &= \begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix} \begin{bmatrix} f^\diamond(\Sigma) & 0 \\ 0 & f^\diamond(\Sigma) \end{bmatrix} \begin{bmatrix} V_1^T & V_2^T \\ -V_2^T & V_1^T \end{bmatrix} \\ &= \begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix} \begin{bmatrix} f(\Sigma) & 0 \\ 0 & f(\Sigma) \end{bmatrix} \begin{bmatrix} V_1^T & V_2^T \\ -V_2^T & V_1^T \end{bmatrix} \\ &= \begin{bmatrix} U_1 f(\Sigma) V_1^T + U_2 f(\Sigma) V_2^T & -U_2 f(\Sigma) V_1^T + U_1 f(\Sigma) V_2^T \\ U_2 f(\Sigma) V_1^T - U_1 f(\Sigma) V_2^T & U_1 f(\Sigma) V_1^T + U_2 f(\Sigma) V_2^T \end{bmatrix}. \end{aligned}$$

Therefore, $\Phi(f^\diamond(A)) = f^\diamond(\Phi(A))$.   □

As a consequence, if a class of $n \times n$ complex matrices $\mathscr{A}$ is invariant under a generalized matrix function $f^\diamond : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$, then the corresponding class of $2n \times 2n$ real matrices $\Phi(\mathscr{A})$ is invariant under the generalized matrix function $f^\diamond : \mathbb{R}^{2n \times 2n} \to \mathbb{R}^{2n \times 2n}$ induced by the same scalar function f; and conversely.

Theorem 10 may be useful in case it is desirable to avoid complex arithmetic in the computation of GMFs of complex matrices.

5 Invariant matrix cones

In Theorem 9 we stated conditions under which the cone of nonnegative matrices in $\mathbb{R}^{m \times n}$ is preserved under GMFs. In this section we consider the invariance of a different type of matrix cone under GMFs.

Let $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ be two fixed unitary matrices. The set $\mathscr{S}_{U,V}$ of all $m \times n$ complex matrices of the form

$$A = U \Sigma V^H, \quad \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p) \in \mathbb{R}^{m \times n}, \quad p = \min\{m, n\},$$

where the σ_i are arbitrary real numbers satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$, is a closed convex cone, i.e., it is closed under nonnegative linear combinations and is closed in $\mathbb{C}^{m \times n}$ under the Euclidean topology. Its interior is the set of all matrices $A \in \mathscr{S}_{U,V}$ with $\mathrm{rank}(A) = p$.

It is obvious that if $f : \mathbb{R} \to \mathbb{R}$ is any nonnegative function which is non-increasing for $x > 0$, then for any choice of U and V the cone $\mathscr{S}_{U,V}$ is invariant under the action of the induced GMF, $f^\diamond$. If $f(x) > 0$ for $x > 0$, then $f^\diamond$ maps the interior of the cone to itself. Hence, if matrices in $\mathscr{S}_{U,V}$ exhibit a certain structure, this structure will be preserved by such GMFs.

As an illustration, consider the case where U and V are discrete Fourier transform matrices of order m and n, respectively. When $m = n$ the set $\mathscr{S}_{U,V}$ is just the positive cone of Hermitian positive semidefinite circulants, and this cone is invariant under any nonnegative GMF; if we restrict our attention to those scalar functions f such that $f(0) = 0$, the induced GMF is actually a standard matrix function. But what happens if we allow $m \neq n$? The following result reveals the structure of complex rectangular matrices of the form $A = U \Sigma V^H$ where $U = U_m$ and $V = U_n$ are DFT matrices of (possibly) different size.

Theorem 11. *Let U_m be the $m \times m$ DFT matrix, let U_n be the $n \times n$ DFT matrix, and let $A \in \mathscr{S}_{U_m, U_n}$. Then A has the following structure:*

$$A = \begin{bmatrix} a & u^H \\ v & B \end{bmatrix},\tag{5}$$

where $a \in \mathbb{R}$, u and v are centrohermitian vectors, and $B \in \mathbb{C}^{(m-1) \times (n-1)}$ is a centrohermitian matrix. In other terms, $u_i = \overline{u_{n-i}}$, $\quad i = 1, \ldots, n-1; v_j = \overline{v_{m-j}}$, $\quad j = 1, \ldots, m-1; b_{tk} = \overline{b}_{m-t, n-k}$, $\quad 1 \le t \le m-1$, $\quad 1 \le k \le n-1$.

Proof. Consider the $m \times m$ circulant permutation matrix

$$P_m = \begin{bmatrix} 1 & & & & \\ & & & & 1 \\ & & & 1 & \\ & & \ddots & & \\ & 1 & & & \end{bmatrix}.$$

Because $\omega^p = \overline{\omega^{km-p}}$, we have that

$$U_m P_m = \frac{1}{\sqrt{m}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{m-1} & \omega^{m-2} & \cdots & \omega \\ 1 & \omega^{2m-2} & \omega^{2m-4} & \cdots & \omega^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(m-1)^2} & \omega^{(m-1)(m-2)} & \cdots & \omega^{(m-1)} \end{bmatrix} = U_m^H.$$

Hence, $U_m P_m U_m = I_m$. Similarly, we have $U_n P_n U_n = I_n$. Therefore, $U_m P_m U_m \Sigma = \Sigma U_n P_n U_n$. It follows that $P_m A = P_m U_m \Sigma U_n^H = U_m^H \Sigma U_n P_n = \overline{U_m \Sigma U_n^H} P_n$, because DFT matrices are symmetric. Thus, we have shown that $P_m A = \overline{A} P_n$.

Partition A into $\begin{bmatrix} a & u^H \\ v & B \end{bmatrix}$, P_m into $\begin{bmatrix} 1 & 0 \\ 0 & \hat{P}_{m-1} \end{bmatrix}$, and P_n into $\begin{bmatrix} 1 & 0 \\ 0 & \hat{P}_{n-1} \end{bmatrix}$. Then

$$\begin{bmatrix} 1 & 0 \\ 0 & \hat{P}_{m-1} \end{bmatrix} \begin{bmatrix} a & u^H \\ v & B \end{bmatrix} = \overline{\begin{bmatrix} a & u^H \\ v & B \end{bmatrix}} \begin{bmatrix} 1 & 0 \\ 0 & \hat{P}_{n-1} \end{bmatrix}$$

and therefore

$$\begin{bmatrix} a & u^H \\ \hat{P}_{m-1} v & \hat{P}_{m-1} B \end{bmatrix} = \overline{\begin{bmatrix} a & u^H \hat{P}_{n-1} \\ v & B \hat{P}_{n-1} \end{bmatrix}}.$$

Thus, a is a real number, u and v are centrohermitian vectors, and B is a centrohermitian matrix [14]. □

As a consequence, any matrix $M = f^\diamond(A)$ with $A \in \mathscr{S}_{U_m, U_n}$ where U_m and U_n are DFT matrices and f is nonnegative and nonincreasing for $x > 0$ has the same structure (5) as A.

6 Conclusions and open questions

In this paper we have identified several types of matrix properties that are preserved under generalized matrix functions. It is interesting to observe that standard and generalized matrix functions behave rather differently from this point of view. There are a few properties that are preserved by both types of matrix function: for instance, being normal or circulant. On the other hand, being triangular or diagonalizable is always preserved by standard matrix functions, but not (in general) by generalized ones. Conversely, Theorem 3 provides several examples of properties that are preserved by generalized matrix functions but not necessarily by standard ones: for instance, $f(A)$ is not generally Hamiltonian when A is Hamiltonian, whereas $f^\diamond(A)$ is always such. We do not claim to have exhausted the list of matrix properties that are preserved by GMFs. Many interesting

open questions remain. For instance, whenever a certain property is not preserved in general, one may ask for conditions on the scalar function f under which the property in question is preserved. In this paper we have seen some instances of this (see Theorems 8-9), but many more such examples are sure to exist.

Finally, although the notion of generalized matrix function applies also to non-square matrices, most of the properties studied in this paper pertain to square matrices only. This is likely a consequence of the fact that square matrices admit a much richer spectral theory and are far more important in applications in other areas of mathematics and science. Nevertheless, it would be interesting to carry out a more complete investigation of structural properties of rectangular matrices and their invariance (or lack thereof) under GMFs.

References

[1]   F. Andersson, M. Carlsson, and K.-M. Perfekt, Operator-Lipschitz estimates for the singular value functional calculus, Proc. Amer. Math. Soc., 144 (2016), pp. 1867–1875.

[2]   F. Arrigo and M. Benzi, Edge modification criteria for enhancing the communicability of digraphs, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 443–468.

[3]   F. Arrigo, M. Benzi, and C. Fenu, Computation of generalized matrix functions, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 836–860.

[4]   B. Arslan, V. Noferini, and F. Tisseur, The structured condition number of a differentiable map between matrix manifolds, with applications, MIMS EPrint 2017.36, Manchester Institute for Mathematical Sciences, 2017.

[5]   J. Aurentz, A. Austin, M. Benzi, and V. Kalantzis, Stable computation of generalized matrix functions via polynomial interpolation, Preprint, 2018.

[6]   P. J. Davis, Circulant Matrices, Wiley, New York, NY, 1979.

[7]   N. Del Buono, L. Lopez, and R. Peluso, Computation of the exponential of large sparse skew-symmetric matrices, SIAM J. Sci. Comput., 27 (2005), pp. 278–293.

[8]   N. Del Buono, L. Lopez and T. Politi, Computation of functions of Hamiltonian and skew-symmetric matrices, Math. Comp. Simul., 79 (2008), pp. 1284–1297.

[9]   J. B. Hawkins and A. Ben-Israel, On generalized matrix functions, Linear and Multilinear Algebra, 1 (1973), pp. 163–171.

[10] N. J. Higham, Functions of Matrices. Theory and Computation, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.

[11] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur, Functions preserving matrix groups and iterations for the matrix square root, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 849–877.

[12] R. D. Hill, R. G. Bates, and S. R. Waters, On perhermitian matrices, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 173–179.

[13] R. A. Horn and C. R. Johnson, Matrix Analysis. Second Edition, Cambridge University Press, 2013.

[14] A. Lee, Centrohermitian and skew-centrohermitian matrices, Linear Algebra Appl., 29 (1980), pp. 205–210.

[15] V. Noferini, A formula for the Fréchet derivative of a generalized matrix function, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 434–457.